STAT 401 Chapter 15

Zepu Zhang
December 3, 2010

The goal of "studies" concerned here is to establish causality. (Recall in regression studies we emphasized that a good model fitting does not establish causality between the predictors and the response.) For example, researchers conduct experiments with a new drug to determine whether the drug has an intended effect. Suppose the experiments are conducted on mice. Some mice take the drug, some don't. The effect of interest will be measured. There are several questions to answer here.

1. Is there significant association between the drug and the effect, say mice that received the drug on average show higher values in the effect of interest?

2. Does the association suggest causality? That is, the increase in the effect is caused by taking the drug, rather than by an external factor.

3. Quantify the causality, if any? For example, the researchers may not only expect the drug to make an impact, but also an impact of certain magnitude. Does the experiment suggest this magnitude is reached?

Whether one can conclude causality largely depends on the design of the experiment. In a properly designed experiment, one can rule out the possibility that the observed effect is caused by an external factor.

Caution: do not jump to concluding causality upon observing association. Check whether the study is properly designed so to facilitate such conclusion.

Example   **quick bread volume**, page 644.
Response: bread volume.
Explanatory variable (factor): baking temperature, 4 levels.
Subjects (experimental units): 20 packages of mix.
Design: assign each of the 4 temperatures to 5 randomly chosen packages. (Completely randomized design.)

Example   **teaching effectiveness**, page 645.
Response: teaching effectiveness (measured in some way) in the year following a teaching seminar.
Explanatory factor: whether the faculty attended the seminar.

Design: each faculty self-decided whether to attend the seminar.

A glossary

factor : explanatory variable. Qualitative or quantitative. If quantitative, it takes several (say 3) values regularly spaced.

factor level : values of the "factor".

treatment : synonym of "factor level".

experimental unit : objects or entities to which treatments are applied.

It is not always obvious what the experimental units are. See (15.3), page 652.

randomization : the procedure to assign treatments to experimental units in a random manner.

Example   How to randomize in the bread volume example:
(1) Create 20 random numbers; assign them in sequence to the 20 packages of mix.
(2) Order the packages based on their random numbers.
(3) Assign treatment 1 (temperature level 1) to the first 5 packages (in the ordered list), treatment 2 to the second 5 packages, and so on.

Randomization is the key to establishing causality. (pages 653–654) "Randomization tends to average out between the treatments whatever systematic effects may be present, apparent or hidden, so that comparison between treatments measure only the pure treatment effects. Thus, randomization tends to eliminate the influence of extraneous factors not under direct control of the experimenter and thereby precludes the presence of selection bias."

experimental studies (15.1) on page 644. In these studies, randomization is used to assign treatments to experimental units, and cause-and-effect relationships between the experimental factors and the outcome or response can be established.

complete replicate If each treatment (or treatment combination) is used exactly n times, we say there are n "complete replicate". In the bread volume example there are 5 complete replicates.

completely randomized design In such a design, the assignment of treatment to each experimental unit is random.

(Probably this can not be defined really rigorously. For example, in the bread volume example the assignment is not "completely" random, because we still stipulate that 5 (rather than a random number of) packages will get treatment 1, 5 get treatment 2, and so on.)

observational studies (15.2) on page 645. In these studies, the association between factor levels and each subject is not "assigned" by the researcher (it is rather by self-choice, or natural occurrence, e.g. disease, or whatever). One "takes whatever there is". "Randomization" is not possible. Of course the researcher can choose the subjects randomly, or by whatever procedure she likes (but the factor level "comes with" the subjects; it is not "assigned" by the researcher).

In these studies, causality is not easy to establish. To do that, there needs to be convincing arguments that the effect can't be a result of external factors.

In the teaching effectiveness example, whether a certain faculty attends the seminar is not assigned by the researcher (who studies relation between teaching effectiveness and attending the seminar).

control : a reference group of experimental units that receive reference level of treatments (e.g. go through the same procedure but do nothing real; taking placebo). They are used for comparison and for "demonstrating" the treatment effects by contrast.

Example   Vitamin C, page 643.
Response: having cold.
Factor: whether to receive a certain vitamin C tablet.
Subjects: 868 children.

blocking pages 655–657, 661–662. If the experimental units are heterogeneous (such as male and female) but are composed of a small number of homogeneous subgroups (such as a subgroup of male and a subgroup of female), the experimental units can be divided into homogeneous subgroups (blocks), and randomization is then done within each block. This way, the treatment effect within each block is not confounded by gender, thus causality can be established.

Note, this is what we do when we do not intend to examine gender effects, which may well exist. We intend

to examine some other effect but want to avoid interference by gender effect, therefore we examine with same-gender groups such that none of the variation in the response is caused by variation in gender. Blocking reduces variance and increases precision, because within a block the variation due to gender is prevented.

factorial design Say there are 2 factors. Factor A has 3 levels and factor B has 4 levels. There are 12 treatment combinations in total. If we randomly assign these 12 treatments to the experimental units, it's called a "completely randomized factorial design".

Example Bread volume, page 659. The model, and test for treatment effect.

Note the alternative H_a: not all β_k equal zero. This necessarily implies that the β_k are not all equal if they are all nonzero. The reason is that any "common" effect would have been picked up by the intercept.

Example Suppose we wanted to study whether 4 wine-making recipes make a difference in the taste of the resultant wine. We made 25 bottles of wine using each of the 4 recipes. Each of the 100 bottles of wine was tasted by a panel of 6 tasters and received an overall score. What are the treatments? What are the experimental units?

Example Suppose we wanted to compare the performance of 2 teaching methods. Of 10 classes, we randomly picked 5 to use method 'A'; the remaining 5 classes used method 'B'. Average exam score of each class was recorded.
(1) What are the treatments?
(2) What are the experimental units?
(3) Is this experimental randomized?
(4) How would you do the randomization?